The Value of Imprecise Prediction

Elliott-Graves, Alkistis

2020

Elliott-Graves , A 2020 , ' The Value of Imprecise Prediction ' , Philosophy, theory, and practice in biology , vol. 12 , no. 4 . https://doi.org/10.3998/ptpbio.16039257.0012.004

http://hdl.handle.net/10138/317304
https://doi.org/10.3998/ptpbio.16039257.0012.004

The Value of Imprecise Prediction

Alkistis Elliott-Graves*

The traditional philosophy of science approach to prediction leaves little room for appreciating the value and potential of imprecise predictions. At best, they are considered a stepping stone to more precise predictions, while at worst they are viewed as detracting from the scientific quality of a discipline. The aim of this paper is to show that imprecise predictions are undervalued in philosophy of science. I review the conceptions of imprecise predictions and the main criticisms levelled against them: (i) that they cannot aid in model selection and improvement, and (ii) that they cannot support effective interventions in practical decision making. I will argue against both criticisms, showing that imprecise predictions have a circumscribed but important and legitimate place in the study of complex, heterogeneous systems. The argument is illustrated and supported by an example from conservation biology, where imprecise models were instrumental in saving the kōkako from extinction.

Keywords

prediction • uncertainty • precision • ecological modelling

1 Introduction

Prediction is important because it constitutes a fundamental facet of scientific practice and an integral feature of the evaluation of scientific theories across many scientific disciplines. However, the notion of prediction in philosophy of science and scientific practice is defined quite narrowly; useful predictions are *precise*, *risky*, and *novel* (Hitchcock and Sober 2004; Lipton 2008; Barrett and Stanford 2006). These characteristics are thought to be crucial when predictions are used to test or confirm theories, their traditional role in philosophy of science (Douglas and Magnus 2013; Lipton 2008; Douglas 2009). Predictions that lack these characteristics consequently receive much less attention, and predominantly negative at that.

One such group is *imprecise predictions*.[1] These are predictions of the existence of a phenomenon, effect or change, without a precise specification of its extent or magnitude. They can take

[1] Sometimes, imprecise predictions are referred to as 'generic' (e.g., in economics, see Rosenberg 1989) or 'qualitative' (e.g., in economics and ecology, see Orzack and Sober 1993; Gonzalez 2015; Dambacher et al. 2003). As we shall see, many of these terms are quite loaded. The most important of these implications is that they tend

*Center for Philosophy of Social Science, Department of Practical Philosophy, P.O. Box 24 (Unioninkatu 40A), FIN-00014 University of Helsinki, Finland, alkistis.elliott-graves@helsinki.fi

Received 20 September 2019; Accepted 1 December 2019
doi:10.3998/ptpbio.16039257.0012.004

the form of trends, ranges or imprecisely probabilistic statements. They occur in a number of sciences, especially those that investigate complex, heterogeneous systems, such as economics, ecology and climate science. For example, imprecise predictions in economics are usually described as predictions **of the existence of a phenomenon**, process, or entity, that do not include any additional specific detail. They "identify the **direction in which changes move**, without however, identifying the **magnitude** of these directions" (Rosenberg 1989). A standard example of an imprecise predictive statement in economics is: 'an increase in tax rate will decrease the firm's output' (without specification of how much the output will decrease) (Rosenberg 1989, 53). In ecology, imprecise predictions are the outputs of mathematical models, such as optimality models, loop analysis, and fuzzy interaction webs (Justus 2006; Dambacher, Li, and Rossignol 2003; Ramsey et al. 2012; Levins 1974; Puccia and Levins 1986). The predictions are trends (i.e., the population will increase/decrease given x, y, z) or ranges (the population will increase by more than x%; the population will decrease by between y and z%). Climate scientists use models to make imprecise predictions about climate variables, such as temperature or precipitation. For example, the Intergovernmental Panel on Climate Change (IPCC)'s fifth report (2013) included the prediction that it is likely (i.e., 60–100% chance) that by 2100 the mean global temperature would rise by between 2.67–4.87°C.

The prevailing attitude towards imprecise predictions is to ignore, dismiss, or criticize them. Philosophers have tended to view them as mathematically unsound, opaque and untrustworthy. For example, in a recent book on economic predictions, Gonzales (2015) defines imprecise predictions as follows:

> 1) The predictions do not follow, in principle, clear rules, because they commonly rely on an intuitive point of view; 2) the subjects … state the future events based on their own experience; 3) the prediction does not detail explicitly how the available information is incorporated into prediction. (58)

At best, they are seen as a preliminary step, to be further refined by more precise quantitative approaches. A famous example of this position comes from Rosenberg's attack of generic predictions in economics:

> Generic predictions are predictions **of the existence of a phenomenon**, process, or entity, as opposed to specific predictions about its detailed character …. This lack of specificity is a weakness in generic theories. … Generic prediction is something, it is a start, but it is not enough. And economists should not be satisfied with it. (1989, 53–55)

Similarly, Orzack and Sober (1993) criticized imprecise (here termed 'qualitative') models in ecology as "not mathematical" (538), while "the idea of qualitative modeling has hindered the development of an unbiased assessment of the truth" (543). They argued that grounds for accepting qualitative predictions are often left unstated and thus inherently suspect:

> The most important defect in qualitative testing … is that it fails to allow one to answer the most important question about a particular model: How well does that model explain the data? Qualitative testing may show some models are incompatible with data, but only quantitative testing of quantitative models can determine what one if any sufficiently explains the data. (542)

to be viewed dichotomously and one side of the dichotomy is thought to lack certain other characteristics (e.g., mathematics). In order to avoid these semantic quibbles, I will be using the term 'imprecise' to refer to predictions that have a certain level of imprecision (though in the spirit of accurate representation, I will retain the other terminology when outlining existing accounts).

For many scientists, reliance on imprecise predictions is seen as a cause of discord within a discipline, or even contributing to the demarcation of a field as 'soft' or 'unscientific' (Peters 1991; Valéry, Fritz, and Lefeuvre 2013; Lipsey 2001; Houlahan et al. 2017; discussion in Winther 2011). For example, Houlahan et al. (2017) criticize their own field (ecology) as not being sufficiently scientific, because it does not pay enough attention to quantitative prediction:

> The lack of emphasis on prediction has resulted in a discipline that tests qualitative, imprecise hypotheses with little concern for whether the results are generalizable beyond where and when the data were collected." (1)[2]

They add that:

> A new commitment to prediction in ecology would lead to, … more mature (i.e., quantitative) hypotheses, prioritization of modeling techniques that are more appropriate for prediction … and, ultimately, advancement towards a more general understanding of the natural world. (1)

A less extreme version of this view is echoed implicitly when scientists argue that a particular science needs to improve its predictive power and suggesting that this can be achieved by increasing the precision of its predictions (Colander et al. 2009; Gurevitch et al. 2011; Hayes and Barry 2007; Evans, Norris, and Benton 2012; Kolar and Lodge 2002). Further evidence of implicit distrust comes from the small number of papers that explicitly advocate the use of models that yield imprecise predictions, and from the ways in which scientists express themselves when they do use them. In many cases, the tone of a paper that produces imprecise predictions is apologetic, highlighting the novelty of the concept or theory being used, or the immature state of the science (see for example Loiselle et al. 2000).

The aim of this paper is to defend a certain subset of imprecise predictions. I start by examining the existing conceptions of imprecise predictions (including qualitative and generic) and classify them in terms of their level of precision (section 2). I then outline the two main criticisms that can be levelled against them: (i) that they are insufficiently risky and therefore cannot be used in model selection or improvement and (ii) that they do not support effective interventions in practical decision making (section 3). In section 4, I defend imprecise predictions by showing both that they can be risky and effectively support interventions. In section 5, I will examine a residual concern, namely that imprecise predictions are only useful in a very specific, non-ideal set of circumstances, hence it is still better to aim for precision in our predictions, whenever possible. I will argue, following Levins (1966), that the reasons we cannot reach this ideal lie in the systems themselves, rather than the methods scientists use to investigate them. Thus, imprecise predictions have a limited but important role in scientific practice.

2 Defining Imprecise Predictions

Defining imprecise predictions is not straightforward. In fields such as Business, Management and some social sciences, imprecise predictions are the results of qualitative research methods such as interviews, questionnaires and the construction of narratives, i.e., methods that do not involve quantification (Patton 2014). Here, predictions are interpretations of this information by experts. Experts gather all the information they deem relevant and make an 'educated guess' about the existence of a phenomenon in the future, or the direction in which it will change.

[2] I should note that *here*, 'hypothesis' refers to outputs of models and experiments, hence it is synonymous with 'prediction'.

Examples include the extrapolation of case studies from one area to another, such as the prediction (by experts) that the programme for improving infant nutrition in Tamil Nadu by educating mothers would also improve infant nutrition in Bangladesh (Cartwright 2012).[3] As these are cases where there is no discernible and systematic method that gives rise to the prediction, this is one category of imprecise predictions I will *not* be defending. Instead, I am interested in imprecise predictions that appear in predominantly quantitative disciplines, such as ecology, economics, and climate change (see examples in Introduction). These predictions are the outputs of mathematical models, hence *imprecise model predictions*. They deserve to be defended because they are seen as being in direct competition with highly precise predictions and come off much worse, as they are tarred with the same brush as the 'expert intuition' predictions mentioned above.

In order to define imprecise predictions, we must first take a closer look at the notion of *precision*, a concept that is itself difficult to define, as it also is used in a number of different ways (Matthewson and Weisberg 2009). First, precision in a model can refer to the parameters in the model (parameter precision) or the output of the model (output precision) (Matthewson and Weisberg 2009). In the context of imprecise model predictions, the output of the model is the most relevant, as predictions are the outputs of models. Yet there are links between parameter and output precision, namely that imprecise specifications of parameters *tend* to produce imprecisely specified outputs. Of course, this is not always the case; it is possible for models with finely specified parameters to produce imprecise predictions (e.g., interval rather than point predictions, or families of models that produce precise but different outputs, that are expressed imprecisely when amalgamated). Nonetheless, as parameter imprecision *can* affect output imprecision, we are interested in the former insofar as it affects the latter.

Second, is precision dichotomous or a matter of degree? Orzack and Sober, in their 1993 criticism of Levins, seem to suggest that it is dichotomous, as they define a model as precise if it 'generates point predictions for output parameters' and imprecise if it does not (534). However, I think it is much more useful to think of precision as a matter of degree. Predicting that an effect will increase rather than decrease is less precise than predicting that it will increase by more than 3%, which is in turn less precise than predicting that it will increase by 3.528%.

Third, why are a number of models imprecise? We can think of imprecision as a way of representing *uncertainty* in our models. In general terms, uncertainty refers to an epistemic limitation, i.e., lack of knowledge about the accuracy of a claim or method. Uncertainty in models can pertain to model inputs, i.e., uncertainty about the types of parameters that should be included in the model and/or uncertainty about the values these parameters should take (Parker 2010; Parker and Risbey 2015). Often, parameter imprecision is inversely related to uncertainty as parameters in a model are defined more precisely as our uncertainty about them decreases (Matthewson and Weisberg 2009).[4] More importantly, uncertainty can refer to model outputs. Here, output imprecision is a reflection of our uncertainty regarding the effect we are predicting (Smith and Stern 2011; Regan, Colyvan, and Burgman 2002; Parker 2010). This uncertainty

[3]In this case, the prediction failed, because the experts did not consider the relevant differences between Tamil Nadu and Bangladesh, such as the fact that educating mothers would not cause a shift in practice in Bangladesh, because they do not have control over shopping or food distribution within the household.

[4]Consider, for example, the exponential growth model, described by the equation $dN/dt = rN$. This equation describes how a population N grows if it is affected only by the intrinsic growth rate r. The value of r is different for each population, and it is calculated from data on birth rates, death rates and fecundity. However, this data is often patchy or incomplete, so scientists might not be certain about the precise value of r. Thus, a complete set of data will allow scientists to express r precisely to many decimal points, e.g., 1.35862, whereas patchier data sets will be expressed less precisely, e.g., as (1.30 ± 0.01) where r ranges from 1.29 to 1.31. In cases of higher uncertainty, r can be expressed as (1.3 ± 0.1), i.e., the range from 1.2 to 1.4, and so on.

can be expressed in a number of ways, the most common of which are: (i) imprecision in terms of the magnitude of the effect we are predicting, which can take the form of interval probability predictions, range predictions or predictions of trends, and (ii) qualitative indications of confidence levels in the accuracy of the prediction in question (Parker and Risbey 2015).

In fact, according to Parker and Risbey (2015), the Intergovernmental Panel on Climate Change (IPCC)'s classification of uncertainty reports can be understood in terms of how precise they are (2–3):

(a) gives a full probability density function/probability distribution over values of X

(b) gives a range of values of X in which the future value can be expected to fall with a precisely specified probability, such as 0.95

(c) gives a range of values of X in which the future value can be expected to fall with an imprecise or interval probability, such as 0.6–0.9, or with a qualitative level of confidence, e.g., medium

(d) gives a range of values of X that can be considered plausible but indicates that probabilities cannot be assigned

(e) gives an order of magnitude estimate of the future value of X but indicates that more precise estimates are out of reach

(f) indicates that the future value of X will be greater than (or less than) the current value, though by how much is unclear

(g) admits that almost nothing is known about the future value of X

This list is a classification of output uncertainty in decreasing levels of precision, i.e., the more uncertain we are about our model outputs, the more imprecise their values. The first point to note is that it shows that a dichotomous notion of precision is unlikely to be useful for predictions under conditions of uncertainty. If precision were dichotomous, then where would the cut-off be? Does the inclusion of probability automatically make a prediction imprecise? Even if it does, this does not mean that all non-precise predictions are equally imprecise. After all, of the main advantages of introducing probabilities into prediction is so that we can differentiate between different levels of imprecision.

Examining the list itself, how does each level relate to imprecise predictions? Starting at the bottom, (g) can be excluded because it does not involve prediction, imprecise or otherwise. Moving up, (f) is probably the most uncontroversial way to characterize an imprecise prediction. It tells about the quality of the change (more or less, plus or minus, greater or lesser) but nothing about the quantity of the change. An example could be: average global temperature will increase in the next 10 years. The next level (e) is more precise, but only slightly. There is still no quantitative information about the change, just a more precise qualitative modifier. For example, the average global temperature will increase more over the next 10 years than in the last 10 years.

Similarly for (d), even though X is expressed in terms of a range of numerical values, there is no *quantification* of uncertainty. We might be able to give a qualitative statement of our expectation for X coming about, e.g., low, medium, high, but this type of statement is not probabilistic. In fact, this type of statement can be distinguished from a confidence interval (as in level (c)), where the qualitative statement is attached to an (imprecise) probability. Here, the qualitative statement is less precise as it refers to the effect itself, rather than our confidence in the probabilistic estimate of the effect. In addition, it may only make sense within a particular context, for example, the statement "the effect of the rat population on the possum population is *low*" might be accurate in one context (i.e., the rat, possum, kōkako (a type of bird) community in the North Island of New Zealand) but may not be accurate in other contexts, such as the effect of an invasive plant species on a community in Australia, where a similar effect (in terms of reduction of population size) may be considered moderate, or even high in that context.

So far, this classification corresponds to scientists' own classifications: levels (e) and (f) correspond to what many economists count as generic predictions. For example, IS-LM models, historically perhaps the most influential macroeconomic models, are highly imprecise. They represent the demand side of the economy with two equations referring to the real sector (investment and saving) and the monetary sector (liquidity preference and monetary supply) and can be used to make predictions about the effect of fiscal and monetary policies on the equilibrium level of income (increase/decrease) (Vercelli 2000).

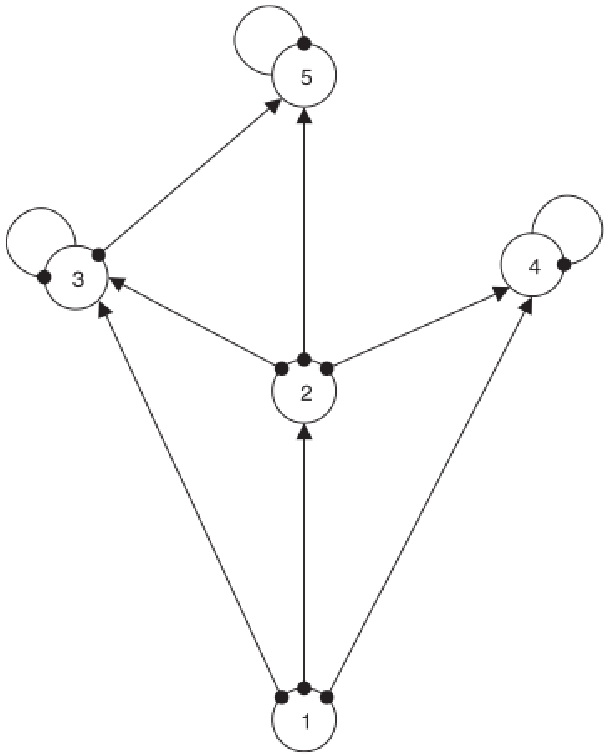
Level (d) is best exemplified by a subset of imprecise predictions in ecology (most others correspond to (c)). An important example of (d) is *loop analysis* (Dambacher, Li, and Rossignol 2003; Justus 2005; Levins 1974; Puccia and Levins 1986). Here, ecologists use signed directed graphs to represent interactions between populations in a community. The graph (and associated community matrix) can be used to identify the *quality* of the effect (positive (1), neutral (0), negative (-1))) of each population on the others and to predict the effects of perturbations of the system, taking into account dynamic qualities of the system such as feedback. For example, in a community of plants, kōkako, rats, possums and stoats, loop analysis can be used to determine all the interactions between each population in the community (see figure 1A) (Ramsey and Veltman 2005). Scientists can use the conjunction of qualitative effects to determine which predator population should be the focus of an intervention so that the prey population is saved from extinction. In this example, loop analysis predicted that an increase in all predator populations would have a negative effect on the kōkako population, but that an increase in the rat population would have a stronger effect than the other two predators, because it would result in a net total of two negative feedback cycles (one from predation and one from competition); see figure 1B.

Moving to the top end of the list, levels (a) and (b) are precise, because they provide *precise probabilistic values* for X. This is in line with scientific practice, where merely containing probabilities is not a reason to think of a prediction as imprecise (Kaplan and Garrick 1981). Introducing probabilities is a way of specifying uncertainty. The important point is that the probabilities themselves are precise, i.e., they are point probabilities. This also goes for (b), where the prediction is for a range of values for X. While a range of values is less precise than a single value, the range still has a precise probabilistic value attached to it.

This leaves level (c), which is the most complicated. Some scientists might classify it as precise merely because it includes numerical values for X. If we do not want it to count as fully precise then we must determine what makes it different from level (b). The key difference is that the probabilities in (c) are themselves imprecise. A probabilistic range, i.e., an interval, is less precise than a point prediction, even if the point prediction is for a range of values. That is, giving a range of probabilities for a range of values reflects more uncertainty than a point prediction for the same range of values. As Parker and Risbey state, this level of imprecision (i.e., a range) can also be equivalently expressed as a qualitative level of confidence for a point probability.

To sum up the discussion in this section, the precise/imprecise distinction should not be viewed as dichotomous (at least in the case of predictions), as model output imprecision comes in degrees. Therefore, both critiques and defences of imprecise predictions should specify the level of imprecision of the prediction in question. In what follows, I will be focusing on imprecise model predictions, which are defined as the outputs of models with a certain level of imprecision.

A. B.



	Fruit/ foliage	Kokako	Ship rats	Possums	Stoats
Fruit/foliage	4	−4	2	0	2
Kokako	−2	3	−2	−1	−1
Ship rats	2	−2	2	0	0
Possums	2	−1	0	1	1
Stoats	0	1	0	−1	1

Figure 1: Loop Analysis: Signed digraph of the kōkako food web (A) and predictions from the adjoint of the qualitative matrix (°A) of interactions from the kokako food web (B). (A) Arrows represent positive causal links and closed circles represent negative causal links. Closed circles starting and ending at the same variable represent self-damping (density-dependent) effects. (B) Values are the net number of feedback cycles (positive and negative) contributing to the response in species i (rows) resulting from a sustained positive input into species j (columns); e.g., the predicted response of a sustained increase in ship rats is predicted to have a negative effect on kokako (i.e., a net total of two negative feedback cycles). Adapted and reprinted with permission from Ramsey and Veltman 2005.

These are the models whose outputs correspond to levels (c)–(f) in the above list. I hope that the discussion so far has shown that this grouping is not arbitrary. It encompasses predictions that are the outputs of *models*, not merely educated guesses, but does not include point predictions or predictions with precise probabilities, both of which count as precise predictions. Moreover, this grouping includes the most prominent imprecise models used by scientists in various fields, and which are subject to the most direct critiques.

Finally, I should note once more that while some of these models and their predictions are labelled 'qualitative' or 'generic' by the scientists who employ them, these terms can be misleading (see for example, Dambacher et al. 2003; Justus 2006; Levins 1974). Labelling a prediction as 'qualitative', might intuitively imply a methodology that involves no quantification, which is not the case in any of the models and predictions I am defending. Thus, it is best to adopt the more neutral terminology of 'precise' (also maximally precise, highly precise) and 'imprecise' (also relatively imprecise, highly imprecise) throughout the discussion. Again, however, I will be using the terms precise and imprecise for the sake of brevity; this is not meant to indicate a strict dichotomy. Imprecise predictions are meant to encompass various levels of imprecision (levels (c)–(f)), while precise predictions are meant to encompass a range of more precise predictions (levels (a)–(b)). In the next two sections, I will examine which critiques can legitimately be levelled against these predictions and provide defences from the critiques.

3 The Case against Imprecise Predictions

If defining imprecise predictions is not straightforward, then identifying their criticisms is even more complicated. The main difficulty is that there are few accounts that actually spell out the criticisms in any detail or provide clear arguments against imprecise predictions. Unfortunately, this does not mean that the critics represent a minority view, rather it shows that the critiques have been widely accepted and entrenched in philosophical and scientific thought, so imprecise predictions are nowadays usually just dismissed without much further thought.[5] Yet this is precisely why imprecise predictions merit a closer look. It is important to identify the strongest criticisms that can be levelled against them and to examine whether they can be overcome.

I will focus on the two most important criticisms of imprecise predictions: (i) that they are insufficiently risky and therefore cannot help us test models or choose between them and (ii) that they cannot support effective interventions. They are the most important because they correspond to the two main functions we expect of any scientific prediction, to be able to test models/theories and to help scientists successfully intervene in real-world situations.

The first criticism addresses the traditional role afforded to prediction. Models or theories are often developed in order to explain phenomena or patterns of data, yet the best way we have to test them is to make predictions based on the model/theory and see if those predictions come out true (Lipton 2005; Douglas 2009). Models/theories that consistently produce accurate predictions are considered successful. Moreover, we can compare the 'track record' of two or more models/theories in order to choose which model to use in a particular context. However, according to Orzack and Sober (1993), if we only have imprecise predictions, we cannot discriminate between two different explanations of a set of data.[6]

They point to the use of optimality models in evolutionary biology, which identify the optimal 'evolutionary behaviour' for a particular context and are used to test the extent to which natural selection rather than other evolutionary forces is the cause of a particular trait (a trait produced by natural selection is considered optimal) (542). For example, they argue that if a model of sex ratios predicted an optimal ratio of 0.95, but an individual produced a sex ratio of 0.6, then this would be cause for concern, and would prompt us to re-examine the model and the data (543). However, if the model predicted the sex ratios less precisely, then the data point of 0.6 would not be detected as a discrepancy and we would not be able to tell if this were a real "lack of optimality" or a "misunderstanding of the biology such that the fact of optimality is not detected" (543). They worry that the predominance of imprecise optimality models "has hindered the development of an unbiased assessment of the truth of one of the most important and influential hypotheses of evolutionary biology" (543).

Why are imprecise predictions not good tests of models' explanations of data? Orzack and Sober do not state this explicitly, but their criticism is based on a deeper issue: the notion that in order for a prediction to count as a true test of a model or theory, it must be sufficiently *risky*. Risky predictions are those that *could* turn out to be false, which is why they are considered good tests. The less obvious a prediction is, given our current knowledge of the model/data, the stronger it is as a test. An example often used in textbooks is Mendeleev's prediction that there were three elements yet to be discovered, with specific physical and chemical properties, such as

[5] The two most important explicit critiques of imprecise predictions are Rosenberg (1989) and Orzack and Sober (1993). For implicit criticism/dismissal of imprecise predictions in ecology and economics see (Colander et al. 2009; Evans et al. 2012; Gurevitch et al. 2011; Hayes and Barry 2007; Houlahan et al. 2017; Kolar and Lodge 2002).

[6] A similar criticism can also be found in Rosenberg's criticism of generic predictions in economics. He argues that Keynesian macroeconomic models (such as the IS-LM) mischaracterized the relationship between unemployment and inflation, which was overlooked because of the imprecision of the predictions(1989, 56).

atomic weights, acidity, specific gravity, etc., which would fit into in to the gaps of the periodic table. This set of predictions is treated as an exemplar of riskiness, because there are many ways in which it could turn out to be false: e.g., there could be fewer or more than three elements, while each element could have different characteristics than the ones predicted.

Precision is one way of increasing the riskiness of a prediction.[7] For example, we could test a model of predation by predicting that the prey population (of, say, bison) will fall from 800 to 500 individuals (e.g., because of a conservation program, such as the introduction of wolves), then rise to 750 individuals (because of the corresponding drop in predator population after the initial drop in prey population). We can then observe the population and collect the relevant data. If the actual prey population does drop to 500 and then rise to 750, we have good reasons to think that the model has latched on to the relevant causal factors of the actual system. If, on the other hand, the population does not eventually rise, or rises to much more or less than 750 individuals (say 900 or 500), then this would mean that there is something wrong with our model. (I am assuming, for the sake of the argument, that we have not made any mistakes in the collection, measurement or interpretation of data.) Moreover, we can use precise predictions to choose between competing models. For example, if we had a second quantitative model which that the prey population would fall to 650 individuals, then we would have a clear reason to favour of the first model, because the first would be more accurate than the second.

Contrast this with a model that predicts only that the prey population will fall, if a predator population is introduced. This is not a completely trivial or risk-free prediction. It could be false if the prey population rose to more than its initial size. However, it would be true whether the population fell to 0, 200, 400 or 750 individuals, so it much easier to confirm. This is problematic because it doesn't provide us with a good enough test of our model. A drop in the prey population to 200 individuals could be because of a different mechanism than a drop to 750 individuals, but the imprecise model and its corresponding prediction would not help us distinguish between the two cases. Nor would it help us to choose between two imprecise models that both predicted a drop in prey population after the introduction of predators.

The second criticism is the worry that scientists need precise predictions in order to accurately anticipate phenomena or changes in real-world systems, so that they can effectively intervene to deal with the situation. The idea is that imprecise model predictions are not always sufficiently informative for effective interventions.[8] There are two ways in which this criticism can be understood. The first is as a practical extension of the first criticism. An imprecise model prediction might point to a correct trend, yet obscure the actual causes of the phenomenon, so an intervention based on that model would be ineffective. Going back to the predation example, the population drop in the prey could be caused by a completely different factor, say a parasite. This would probably cause a much larger drop in the prey population (e.g., to 200) and would not include the subsequent rise (e.g., to 750). As the imprecise prediction does not specify the extent of the population drop, we might fail to recognise the true cause of the drop in the prey population and intervene ineffectively (i.e., by controlling the predators).

The second version of the criticism is even more pragmatic. Even if we latch on to the accurate mechanism causing the population drop, then the lack of precise information could also cause us to intervene ineffectively. For example, a small drop, say to 750 individuals, does not usually require an intervention (the population will often bounce back on its own), whereas a larger drop, e.g., to 200 individuals, can be much more dangerous and merit an intervention if

[7]The other most common characteristic of riskiness is *novelty*. A discussion of novelty is beyond the scope of this paper, but see (Hitchcock and Sober 2004; Douglas and Magnus 2013).

[8]Rosenberg uses this criticism in his attack on Keynesian macroeconomic models, which he believes led to ineffective interventions on the economy (1989, 58–59).

the population is to be conserved. With just the output of the imprecise model, scientists are at risk of not realizing that an intervention is necessary if the population drop is large. Alternatively, they might spend scarce resources to intervene without it being necessary if the population drop is small.

These are the strongest criticisms of imprecise model predictions, because they address the two most important roles any scientific prediction ought to embody. If they cannot be overcome, then there is really no reason to invest time, energy and funding on imprecise predictions.

4 Defending Imprecise Predictions

4.1 The Case of the North Island Kōkako

The North Island kōkako is a bird endemic to New Zealand with a beautiful song (Ramsey and Veltman 2005, henceforth R&V). In 1999 the species was reduced to 400 pairs. The cause of the kōkako decline was predation. This is not surprising, in itself, but there are three main bird predators in New Zealand, the so-called 'unholy trinity': rats, possums and stoats. As resources were limited, it was important to determine which of these predators to focus on primarily, and the extent to which each population needed to be culled. To make matters worse, the policy needed to be researched and implemented before the kōkako numbers dropped further.

R&V used two imprecise models to analyse the dynamics of the kōkako community and to predict the effect of interventions on that community for the kōkako population. The first model was loop analysis (outlined in section 2) and was used to identify all the dynamic relationships between populations in the community, along with their quality (positive, neutral or negative). Adding up the number of positive neutral and negative effects reveals the overall effect of each population on the others. In this case, loop analysis predicted that the rat population had the overall most important effect on the kōkako.

The second model was a 'fuzzy interaction web' (FIW). FIWs take imprecise information on the abundances of populations within an ecological community, i.e., population abundance data that is incomplete or imprecise, and create 'fuzzy sets'. This is because it is practically impossible to determine the exact size of each population, and consequently the precise rates of competition and predation within the community. The scientists express this uncertainty through 'fuzziness' i.e., each element can have partial membership of a set, and/or can belong to multiple sets.

To clarify, let us start with what the scientists do know. They can measure 'tracking rates', i.e., the percentage of traps that are full on any given night. They then determine the 'indexes' of each population i.e., what counts as low, medium or high abundance. This is based on comparing the tracking rates with existing data of past tracking rates (in this case the data stretched back more than 10 years). For example, in this community, when 20% or above of the traps each night are full, this counts as high abundance, whereas an index of below 10% represents low abundance. However, as the scientists at no point know the exact size of any population, this index is an estimate, hence the intermediate percentages have *partial membership* in high, medium and low abundances (see figure 2). In other words, by fuzzifying the membership in each set (high, medium, low), the scientists are incorporating the inherent uncertainties of each population size estimate.

With the indexes in place, the scientists can make predictions about the effect of each population on the other populations in the community. Figure 3 summarises the predictions yielded by the FIW. It shows that controlling all three populations would have a high effect on the kōkako population, however, controlling rats and possums would have a moderate effect on the kōkako population. Given the population indexes mentioned above, a moderate effect is

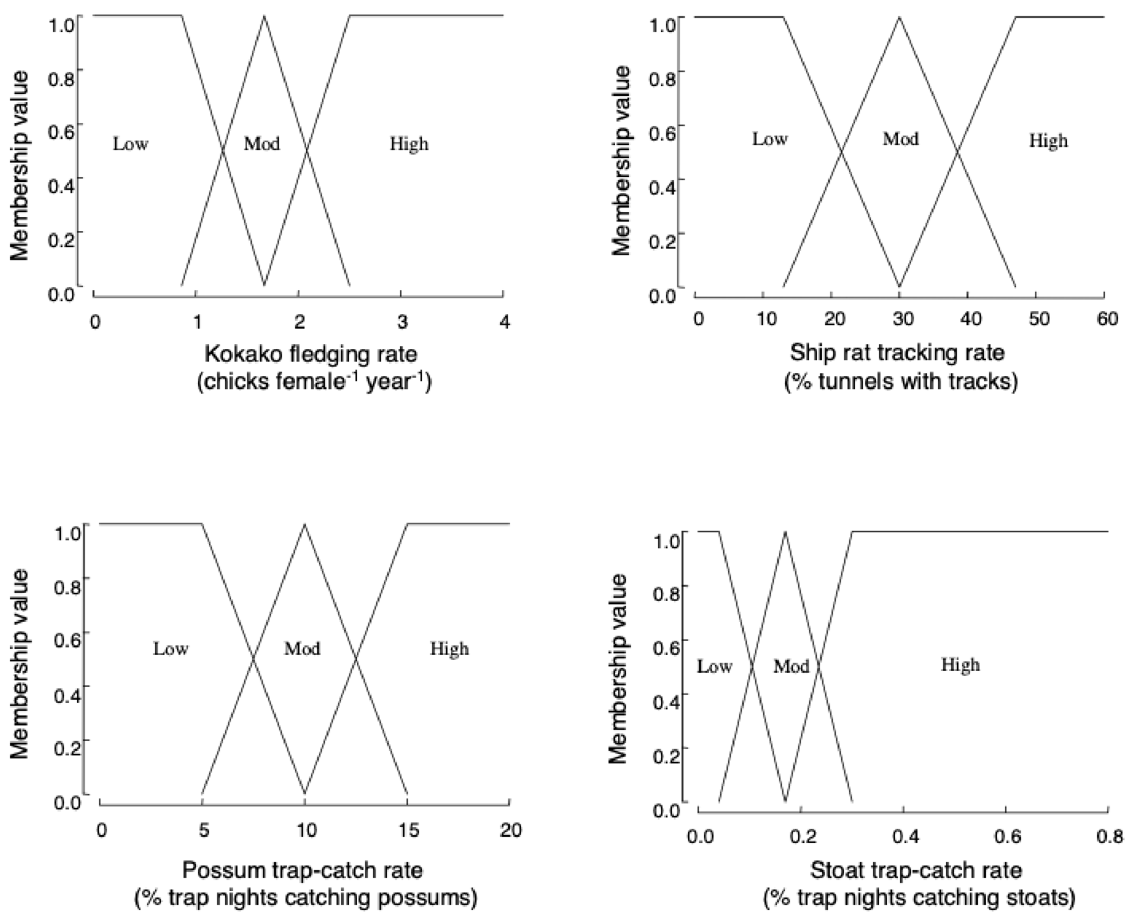


Figure 2: Fuzzy set membership functions for linguistic descriptions of species abundances ('low', 'mod(erate)', 'high') for each of the animal species in the kōkako Fuzzy Interaction Web (reprinted with permission from R&V, supplementary materials)

Species controlled	Population abundance of rats	Population abundance of possums	Population abundance of stoats	Kokako fledging rate
Rats	Low	Mod	Low	Low
Possums	Mod	Low	Mod	Low
Stoats	Mod	Mod	Low	Low
Rats and possums	Low	Low	Mod	Mod
Rats and stoats	Low	Mod	Low	Low
Possums and stoats	Mod	Low	Low	Low
All three	Low	Low	Low	High

Figure 3: Fuzzy Interaction Web Predictions: Imprecise predictions of the magnitude of the effect on the equilibrium kōkako fledgling rate resulting from sustained single and multispecies control of nest predators from the FIW 'trained' model. (Reprinted with permission from R&V.)

sufficient for bringing back the kōkako population to acceptable levels, thus the most efficient intervention should focus on the rats and possums. More specifically, the FIW predicted that the rats need to be kept at a tracking rate of "about 11% or lower" and possums at a tracking rate "below about 10%", so that the kōkako population is maintained at "moderate levels" (R&V, 914).

This example reveals two important points about the value of imprecise model predictions. The first is that both models are genuinely imprecise (loop analysis corresponds to level (d), and FIWs correspond to level (c) in the classification outlined in section 2). At no point do the scientists have precise estimations of each population or even precise estimations of what percentage of the population should be culled. Instead, they can make comparisons between the strength of the effects between the populations of the community and imprecise estimates of tracking rates. That is, they do not need to know how large the populations actually are, they merely need to know to keep culling until the tracking rates are below a certain percentage.

This is actually important practical advantage of imprecise models. A common problem in ecology (and other disciplines) is paucity of data: often, the data we want are difficult to obtain, or even unavailable. For example, the capture-recapture method is perhaps the most widely used sampling method in ecology yet does not work equally well for all species/populations. It is not very useful in cases where capture or recapture is difficult, and it can result in biased samples. For instance, in some fish species, capture methods are biased towards larger fish, skewing demographic information (Pine et al. 2003).

In cases like these, imprecise rather than highly precise models may be superior. There have been studies that compare the success rates of imprecise and highly precise predictions in a number of contexts. Precise models tend to be much more susceptible to poor, biased or patchy data, precisely because of their finely specified variables. For example, (Novak et al. 2011) found that even if there is observational and experimental data on *particular species within the ecosystem* but insufficient data on the *indirect effects* of each interaction on other populations within the ecosystem or other dynamics, the predictive accuracy of highly precise models dwindles rapidly (e.g., to about the level of flipping a coin). In addition, the process of transforming imprecise data to precise variables for highly precise models is generally ill-advised, because while it results in model outputs that are precise numbers, they are often wildly inaccurate (e.g., population in the future is much larger/smaller than the model predicts), or not particularly illuminating (e.g., the sign $(+/-)$ of the number can vary, which means that we cannot be sure that the interaction is positive or negative) (Dambacher, Li, and Rossignol 2003).

The second point directly addresses the issue of *riskiness* and *model choice*. A non-risky prediction is that *all* predator populations have negative effects on the kōkako. This is quite intuitive. Moreover, it would not be falsified if culling any one, two or all three predator populations had the desired effect on the kōkako. However, neither loop analysis nor the FIWs made that prediction. As we saw, in both cases the models predicted that culling all three predators would have a positive effect on the kōkako population, yet they both made the much riskier prediction that this intervention was unnecessary, and that the intervention could focus on just the rat population (loop analysis) or the rat and possum population (FIW).

In fact, as the two models made different predictions, this gives us an opportunity to test each model. This is precisely what the scientists did in their paper, revealing an important flaw in loop analysis that could be overcome with the FIW. This is because in loop analysis, the (strong) negative feedback cycle (between rats and stoats) was counteracted by 2 (weak) positive feedback cycles. As loop analysis cannot distinguish between feedback cycle strengths, the prediction of the effect was lost. In contrast, the FIW examined the effect of each *pair of predators* on the kōkako population. Thus, it predicted that the stoat population is affected by the rat population, hence a reduction in the rat population would lead to a reduction in the stoat population. This means that intervening on the *rat* population renders intervention on the *stoat* population unnecessary, but also controlling the *possum* population would make the intervention more successful.

I should note that ecologists working with imprecise model predictions are remarkably careful in using and testing their models. As stated above, Novak et al. (2011) compared the predictive success of imprecise and highly precise models, showing the advantages and limitations of each. R&V were aware of the limitations of loop analysis, hence used a second model to test the data. Furthermore, they actually examined the possibility of using highly precise models for the kōkako community and showed that the available data on the estimates of interaction strengths did not conform to the parameters required by highly precise models, so that they resulted in wildly inaccurate results (906).

Moving to the second criticism, did these predictions support effective interventions? They did. The case of the kōkako is actually a heartening success story of interventionist conservation. Today, there are about 1600 pairs dispersed over 22 different sites. 10 of these populations have been recovered and an additional 12 have been newly established. The plans continue in the future, with the goal being to increase the population to 3000 pairs by 2025 (New Zealand Department of Conservation 2017).

As it turns out, we also know that the model predictions I have discussed were actually used in the intervention on the kōkako and were instrumental in saving the population from extinction. One of the paper's authors (Veltman), was based at the Science and Research Unit of the New Zealand Department of Conservation, which was in charge of the kōkako conservation project. In addition, other papers she (co)authored include research on direct and indirect effects of pest (rat, stoat, and possum) control on ecological communities (Tompkins and Veltman 2006). This work was part of a larger project within the Department of Conservation, contributing to a paper on imprecise models for control of deer, which are also a pest in New Zealand (Ramsey et al. 2012). In addition, subsequent publications of the Department of Conservation incorporate the research and recommendations from papers by these scientists (see, for example, Brown, Elliott, and Kemp 2015).

To sum up, the case of the north Island kōkako provides us with an example of imprecise model predictions that were both risky and supported effective interventions. Yet the kōkako case required immediate intervention and was constrained by paucity of data. How often do these types of situations come about? In other words, what is the scope of imprecise model predictions in scientific practice? I will address this worry in the next section.

4.2 The Role of Imprecise Model Predictions in Scientific Practice

What exactly is the role of imprecise model predictions in scientific practice? When should scientists prefer models with imprecise predictions over models with precise predictions? My aim in this section is not to claim that imprecise predictions are *always* preferable to maximally precise ones. That claim would be clearly false. However, I am claiming that there are some contexts in which imprecise predictions are *consistently* preferable. The most comprehensive account of the context for imprecise predictions can be found in the work of Richard Levins. A substantial part of his work was geared towards creating the conceptual space and providing support for qualitative analysis and imprecise modelling (Justus 2006). His seminal paper on different strategies for model-building (Levins 1966) can be understood, in part, as providing a theoretical argument for model diversity, which includes the motivation for constructing imprecise models (Weisberg 2006).

This motivation is key to understanding the importance of imprecise predictions for scientific practice. Levins famously pointed out that systems in the natural world are very complex (i.e., made up of many interacting parts), and that this complexity cannot be incorporated in our models, in its entirety. We aim to maximize the quality of our models by making them as *precise*, *realistic* and *general* as possible, yet complexity results in a tradeoff between these desiderata. Only two of the three desiderata can be maximized in each model, thus giving rise to three strategies for model building (each sacrificing one desideratum and maximizing the other two) and three corresponding types of models.

In this context, imprecise modelling, which produces imprecise predictions as its output, is one of three legitimate strategies for model-building (type III). That is, by sacrificing precision, imprecise models are able to maximize the other two desiderata. First, by sacrificing precision instead of *realism* they are said to contain *idealizations of specificity* rather than *idealizations of veracity* (Justus 2006, 659). That is, imprecise models represent systems *veridically*, yet are simplified in the sense that particular properties are represented with low degrees of specificity. For example, the FIW included *all* populations in the community, the resources (fruit and foliage) and *all* the interactions between each population, i.e., not just the effect of each predator on the kōkako. The simplification here was that no population size was represented precisely.

In contrast, maximally precise models (usually type I, sacrificing realism for generality and precision) employ different methods of simplification, i.e., making unrealistic assumptions and/or omitting causal factors altogether. For example, the most common alternative to imprecise models for community interactions is a version of the Lotka-Volterra predation model.[9] This model does not include the resources (fruit and foliage), and simplifies the interactions between each population, i.e., it does not separately identify the interactions between each population in the community (see footnote 7, R&V, 906, and Dambacher et al. 2003 for a full explanation of the details). Not all factors make it into the model and many that do are simplified, decreasing the 'veracity' of the model. The problem is that this type of idealization often mischaracterizes

[9] The version presented in R&V is a multi-species interaction model, described by the equation $dN_i/dt = f_i(N) i = 1, 2, \ldots, s$, where N_i is the density of species i, N is the vector of s species densities $N_1, N_2, \ldots N_s$, and f_i is the function describing the growth rate of species i. The partial derivative of f_i with respect to species j at equilibrium growth rates yields $a_{ij} = \partial f_i / \partial N_j$. Based on this, the scientists construct a community matrix to make predictions based on a sustained increase or decrease of any of the community's members (all the matrix's elements are represented together the term a_{ij}, so any change in one of the community members reflects a change in a_{ij}) (906).

salient features of systems, resulting in inaccurate explanations and/or predictions (Justus 2005, 2006). In the above example, the idealizations of veracity result in the models not being able to predict the effects of large perturbations or non-linear interactions, which are very common in real communities (R&V, 906; Dambacher et al. 2003, 80).

Second, allowing for imprecision results in higher levels of *generality*, i.e., the models are not tied to a particular system, but applicable to many (different) systems. Generalizing results can be very useful, because it allows scientists to make connections (comparisons, contrasts) across systems. In some cases, general models can pave the way to unifying explanations that subsume a large number of phenomena (Weisberg 2006). The problem in complex systems, is that generality is usually either abandoned (type II models) or achieved at the expense of realism (type I models), models that are applicable widely but whose explanations and predictions are often inaccurate (Elliott-Graves 2018). In contrast, by sacrificing precision, ecologists such as Levins were able to construct models that were widely applicable to many different populations in different environments, but also realistic, incorporating many of the salient variables for each population and thus minimizing the risk of inaccuracies.

The tradeoff account provides the basic insight of how complexity[10] causes difficulties studying ecological systems but does not adequately capture the extent of these difficulties. First, many ecological models incorporate complexity (type II models). Second, scientists in other fields that study complex systems, such as physics and chemistry, do not seem to face such extensive tradeoffs (Justus 2005 in Matthewson 2011). A Boeing 747 is a complex entity, yet once we have knowledge of how one Boeing 747 works, we can generalize to other cases and make accurate predictions about a range of flight trajectories, etc. (Matthewson 2011).

The answer lies in another feature that *magnifies* the tradeoff between generality realism and precision. *Causal heterogeneity* occurs when the parts that make up a system are themselves diverse (Elliott-Graves 2018).[11] That is, ecosystems or economies are not like aeroplanes; none is identical to the others. This means that generalising across systems is often difficult or even impossible, as knowledge of what happens in one system does not necessarily apply to other systems, even if they seem similar at first glance. This is precisely what has happened in many cases of simple, precise and general models (type I) in ecology, where a model that works for one system produces inaccurate predictions in other systems. For example, general models of competition, that apply well to animal populations often yield inaccurate predictions for plant populations (Berger et al. 2008). In contrast, imprecise models can incorporate a much larger number of causal factors, without being tied to a particular system, because outputs are specified imprecisely. For example, a loop analysis model for a community of n species will be applicable to all communities of n species, irrespective of their geographical location, and will include all interactions between populations and resources. Thus, it will capture every single relevant interaction and its quality (positive, neutral or negative), though it will not specify its magnitude.

This is especially pertinent in the case of prediction because generalizations *across time* are also often unwarranted (Elliott-Graves 2018). Even if scientists have adequate knowledge of the functioning of a system at a particular time, the existence of dynamics, feedback, threshold effects, etc. mean that this knowledge might not be projectable into the future. For example, current knowledge of how insects deal with the cold during wintertime might not result in

[10] In this context (and literature), complexity is usually understood to mean the following: a system is complex when it has many interacting parts (Elliott-Graves 2018; Levins 1966; Matthewson and Weisberg 2009; Matthewson 2011). Systems are also sometimes described as complex when they exhibit emergent behaviours (Matthewson 2011).

[11] A full examination of causal heterogeneity and all its effects on scientific practice is beyond the scope of this paper but can be found in Elliott-Graves 2018.

accurate predictions of native and invasive insect abundances in the light of climate change (Marshall and Sinclair 2012; Kaunisto, Ferguson, and Sinclair 2016).

To sum up the argument so far, sacrificing realism in complex heterogeneous systems is dangerous, because we need sufficient causal information in order to make accurate predictions. The causal heterogeneity here matters more than the complexity. As stated above, models that sacrifice realism leave out causal factors, and focus on a small subset of them. This is not a problem in complex systems *if* the factors that remain in the model are the relevant causal factors whereas the factors left out are mere details. The problem in *causally heterogeneous* systems is that they differ in terms of causes. Thus, any time a factor is left out, it has a much higher chance of being relevant. In these types of cases, imprecise model predictions have a higher chance of providing accurate predictions, hence should preferred, or at the very least, genuinely considered.

Finally, even if this argument is persuasive within the context described, a critic of imprecise predictions can still maintain that this context is an example of non-ideal or immature science. For example, Rosenberg allowed that imprecise predictions in economics can be useful in certain cases, though he maintained that they should be superseded by more precise predictions whenever and wherever possible. If a science is to be considered fully mature, then scientists should do whatever it takes to get closer to the ideal of maximal precision. The implication is that the difficulties leading to the use of imprecise models are practical, such as lack of data, poor quality data, working under time constraints (all of which apply to the kōkako case). But, the argument goes, with time, scientists will build more extensive data sets and get better knowledge of the systems they are investigating so that maximally precise predictions will be possible and imprecise predictions superfluous. Therefore, scientists should not become complacent and accept that imprecise predictions are good enough, even if they occasionally are useful.

The problem with this criticism is that it is based on a misunderstanding. Complexity and causal heterogeneity are not artefacts of our scientific methods, but intrinsic factors of many systems under investigation. Moreover, they are factors that cannot be overcome, even as science progresses. Even if we acquire more data and gain a better understanding of each system we study, this does not mean that the same causal factors will be relevant in other systems or at other times. This may seem far from our picture of an 'ideal' science, but it does not mean that we can shoehorn these systems into the pre-existing ideal. As we have seen, ignoring or reducing complexity and causal heterogeneity in our models often makes our scientific investigations worse, as the subsequent predictions are inaccurate. Yet we have an alternative option for these cases, namely sacrificing precision with imprecise models. As Levins and others have shown, this is a legitimate option, which is underutilized. It is perhaps this preoccupation with a singular conception of an 'ideal science' that has led us to overlook the contexts in which imprecise predictions are useful and thus undervalue them.

5 Conclusion

Imprecise predictions are undervalued in philosophy of science, as there is some confusion about how they should be defined, and because they are mistakenly thought to be opaque, mathematically suspect or useless. I have shown that a large and important subset of imprecise predictions are not subject to these criticisms and in certain contexts they outperform precise predictions. While most of the examples of this analysis were from ecology, a discipline where imprecise predictions have been used and evaluated more thoroughly, the classification and defence of imprecise predictions can be applied to other disciplines that face similar issues of uncertainty and causal heterogeneity, such as economics and climate science.

Acknowledgments

I would like to thank Kim Sterelny, Ronald Planer, Rachael Brown, and two anonymous reviewers for helpful and constructive comments. This paper is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 796910.

Literature cited

Barrett, Jeff, and P. Kyle Stanford. 2006. "Prediction." In *The Philosophy of Science: An Encyclopedia*, edited by Jessica Pfeifer and Sahotra Sarkar. New York: Routledge.

Berger, Uta, Cyril Piou, Katja Schiffers, and Volker Grimm. 2008. "Competition Among Plants: Concepts, Individual-Based Modelling Approaches, and a Proposal for a Future Research Strategy." *Perspectives in Plant Ecology, Evolution and Systematics* 9 (3–4): 121–35. doi:10.1016/j.ppees.2007.11.002.

Brown, Kerry, John Innes Elliott, and Josh Kemp. 2015. "Ship Rat, Stoat and Possum Control on Mainland New Zealand." New Zealand Department of Conservation.

Cartwright, Nancy. 2012. "Will This Policy Work for You? Predicting Effectiveness Better: How Philosophy Helps." *Philosophy of Science* 79 (5): 973–89. doi:10.1086/668041.

Colander, David, Michael Goldberg, Armin Haas, Katarina Juselius, Alan Kirman, Thomas Lux, and Brigitte Sloth. 2009. "The Financial Crisis and the Systemic Failure of the Economics Profession." *Critical Review*, July. Routledge. doi:10.1080/08913810902934109.

Dambacher, Jeffrey M., Hiram W. Li, and Philippe A. Rossignol. 2003. "Qualitative Predictions in Model Ecosystems." *Ecological Modelling* 161 (1–2): 79–93. doi:10.1016/S0304-3800(02)00295-8.

Douglas, Heather E. 2009. "Reintroducing Prediction to Explanation." *Philosophy of Science* 76 (4): 444–63. doi:10.1086/648111.

Douglas, Heather, and P. D. Magnus. 2013. "State of the Field: Why Novel Prediction Matters." *Studies in History and Philosophy of Science* 44 (4): 580–89. doi:10.1016/j.shpsa.2013.04.001.

Elliott-Graves Alkistis. 2018. "Generality and Causal Interdependence in Ecology." *Philosophy of Science* 85 (1): 1102–1114.

Evans, Matthew R., Ken J. Norris, and Tim G. Benton. 2012. "Predictive Ecology: Systems Approaches." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1586): 163–69. doi:10.1098/rstb.2011.0191.

Gurevitch, J., G. A. Fox, G. M. Wardle, Inderjit, and D. Taub. 2011. "Emergent Insights From the Synthesis of Conceptual Frameworks for Biological Invasions." *Ecology Letters* 14 (4): 407–18. doi:10.1111/j.1461-0248.2011.01594.x.

Hayes, Keith R., and Simon C. Barry. 2007. "Are There Any Consistent Predictors of Invasion Success?" *Biological Invasions* 10 (4): 483–506. doi:10.1007/s10530-007-9146-5.

Hitchcock, Christopher, and Elliott Sober. 2004. "Prediction Versus Accommodation and the Risk of Overfitting." *British Journal for the Philosophy of Science* 55 (1): 1–34. doi:10.1093/bjps/55.1.1.

Houlahan, Jeff E., Shawn T. McKinney, T. Michael Anderson, and Brian J. McGill. 2017. "The Priority of Prediction in Ecological Understanding." *Oikos* 126 (1): 1–7. doi:10.1111/oik.03726.

Justus, James. 2005. "Qualitative Scientific Modeling and Loop Analysis." *Philosophy of Science* 72 (5): 1272–1286. doi:10.1086/508099.

Justus, James. 2006. "Loop Analysis and Qualitative Modeling: Limitations and Merits." *Biology & Philosophy* 21 (5): 647–66. doi:10.1007/s10539-006-9050-x.

Kaplan, Stanley, and B. John Garrick. 1981. "On the Quantitative Definition of Risk." *Risk Analysis* 1 (1): 11–27. doi:10.1111/j.1539-6924.1981.tb01350.x.

Kaunisto, Sirpa, Laura V. Ferguson, and Brent J. Sinclair. 2016. "Can We Predict the Effects of Multiple Stressors on Insects in a Changing Climate?" *Current Opinion in Insect Science* 17 (October): 55–61. doi:10.1016/j.cois.2016.07.001.

Kolar, C. S., and D. M Lodge. 2002. "Ecological Predictions and Risk Assessment for Alien Fishes in North America." *Science* 298 (5596): 1233–36.

Levins, Richard 1966. "The Strategy of Model Building in Population Biology." *American Scientist* 54 (4): 421–31.

Levins, R. 1974. "Qualitative Analysis of Partially Specified Systems." *Annals of the New York Academy of Sciences* 231: 123–38. doi:10.1111/j.1749-6632.1974.tb20562.x.

Lipsey, Richard G. 2001. "Successes and Failures in the Transformation of Economics." *Journal of Economic Methodology* 8 (2): 169–201. doi:10.1080/13501780110047273.

Lipton, Peter. 2005. "Testing Hypotheses: Prediction and Prejudice." *Science* 307 (5707): 219–21. doi:10.1126/science.1103024.

Lipton, Peter. 2008. "Prediction and Prejudice." *International Studies in the Philosophy of Science* 4 (1): 51–65. doi:10.1080/02698599008573345.

Loiselle, Steven, Giuseppe M. Carpaneto, Vincent Hull, Tomas Waller, and Claudio Rossi. 2000. "Feedback Analysis in Reserve Management: Studying Local Myths Using Qualitative Models." *Ecological Modelling* 129 (1): 25–37. doi:10.1016/S0304-3800(00)00217-9.

Marshall, Katie E., and Brent J. Sinclair. 2012. "Threshold Temperatures Mediate the Impact of Reduced Snow Cover on Overwintering Freeze-Tolerant Caterpillars." *Naturwissenschaften* 99 (1): 33–41. doi:10.1007/s00114-011-0866-0.

Matthewson, J., and M. Weisberg. 2009. "The Structure of Tradeoffs in Model Building." *Synthese* 170 (1): 169–90. doi:10.1007/s11229-008-9366-y.

Matthewson, John. 2011. "Trade-Offs in Model-Building: a More Target-Oriented Approach." *Studies in History and Philosophy of Science Part A* 42 (2): 324–33. doi:10.1016/j.shpsa.2010.11.040.

New Zealand Department of Conservation. 2017. "New Zealand's Threatened Species Strategy." Found at https://www.doc.govt.nz/contentassets/97142dceee6b4f0e95a625ab4fa58dfc/threatened-species-strategy-draft.pdf

Novak, Mark, J. Timothy Wootton, Daniel F. Doak, Mark Emmerson, James A. Estes, and M. Timothy Tinker. 2011. "Predicting Community Responses to Perturbations in the Face of Imperfect Knowledge and Network Complexity." *Ecology* 92 (4): 836–46. doi:10.1890/10-1354.1.

Parker, Wendy. 2010. "Predicting Weather and Climate: Uncertainty, Ensembles and Probability." *Studies in History and Philosophy of Modern Physics* 41 (3): 263–72. doi:10.1016/j.shpsb.2010.07.006.

Patton, Michael Quinn. 2014. *Qualitative Research & Evaluation Methods*. SAGE Publications. doi:10.1002/0470013192.bsa514.

Peters, R. H. 1991. *A Critique for Ecology*. Cambridge: Cambridge University Press.

Pine, W. E., K. H Pollock, J. E Hightower, T. J Kwak, and J. A Rice. 2003. "A Review of Tagging Methods for Estimating Fish Population Size and Components of Mortality." *Fisheries* 28 (10): 10–23. doi:10.1577/1548-8446(2003)28[10:AROTMF]2.0.CO;2.

Puccia, Charles J., and Richard Levins. 1986. *Qualitative Modeling of Complex Systems*. Cambridge, MA: Harvard University Press. doi:10.4159/harvard.9780674435070.

Ramsey, David S. L., and Clare J. Veltman. 2005. "Predicting the Effects of Perturbations on Ecological Communities: What Can Qualitative Models Offer?" *Journal of Animal Ecology* 74 (5): 905–16. doi:10.1111/j.1365-2656.2005.00986.x.

Ramsey, David S. L., David M. Forsyth, Clare J. Veltman, Simon J. Nicol, Charles R. Todd, Robert B. Allen, Will J. Allen, et al. 2012. "An Approximate Bayesian Algorithm for Training Fuzzy Cognitive Map Models of Forest Responses to Deer Control in a New Zealand Adaptive Management Experiment." *Ecological Modelling* 240 (August): 93–104. doi:10.1016/j.ecolmodel.2012.04.022.

Regan, Helen M., Mark Colyvan, and Mark A. Burgman. 2002. "A Taxonomy and Treatment of Uncertainty for Ecology and Conservation Biology." *Ecological Applications* 12 (2): 618–28. doi:10.1890/1051-0761(2002)012[0618:ATATOU]2.0.CO;2.

Rosenberg, A. 1989. "Are Generic Predictions Enough?" *Philosophy of Economics*. doi:10.2307/20012198.

Smith, Leonard A., and Nicholas Stern. 2011. "Uncertainty in Science and Its Role in Climate Policy." *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 369 (1956): 4818–41. doi:10.1098/rsta.2011.0149.

Tompkins, Daniel M., and Clare J. Veltman. 2006. "Unexpected Consequences of Vertebrate Pest Control: Predictions From a Four-Species Community Model." *Ecological Applications* 16 (3): 1050–61.

Valéry, Loïc, Hervé Fritz, and Jean-Claude Lefeuvre. 2013. "Another Call for the End of Invasion Biology." *Oikos* 122 (8): 1143–46. doi:10.1111/j.1600-0706.2013.00445.x.

Vercelli, Alessandro. 2000. "The Evolution of IS-LM Models: Empirical Evidence and Theoretical Presuppositions." In *Macroeconomics and the Real World*, 25–42. Oxford: Oxford University Press.

Weisberg, M. 2006. "Forty Years of 'The Strategy': Levins on Model Building and Idealization." *Biology & Philosophy*. doi:10.1007/s10539-006-9051-9.

Winther, Rasmus Grønfeldt. 2011. "Prediction in Selectionist Evolutionary Theory." *Philosophy of Science* 76 (5): 889–901. doi:10.1086/605824.

ISSN 2475-3025